

日 本 国 特 許 庁
JAPAN PATENT OFFICE

J1002 U.S. PTO
10/045512
11/07/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2000年11月29日

出 願 番 号

Application Number:

特願2000-363394

出 願 人

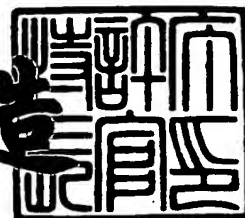
Applicant(s):

松下電器産業株式会社

2001年 5月11日

特 許 庁 長 官
Commissioner,
Japan Patent Office

及 川 耕 造



出証番号 出証特2001-3039113

【書類名】 特許願

【整理番号】 2900825591

【提出日】 平成12年11月29日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 3/00

【発明者】

【住所又は居所】 神奈川県横浜市港北区綱島東四丁目3番1号 松下通信工業株式会社内

【氏名】 磯野 敏幸

【発明者】

【住所又は居所】 神奈川県横浜市港北区綱島東四丁目3番1号 松下通信工業株式会社内

【氏名】 西村 洋文

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100072604

【弁理士】

【氏名又は名称】 有我 軍一郎

【電話番号】 03-3370-2470

【手数料の表示】

【予納台帳番号】 006529

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9908698

特 2 0 0 0 - 3 6 3 3 9 4

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声合成方法およびその装置

【特許請求の範囲】

【請求項 1】 入力するテキストデータを録音音声で再生するときには、前記録音音声をそのまま出力し、波形重畳法で作成した合成音声で出力するときには、前記録音音声を録音した環境と同じ残響特性をもつ残響特性を付与して出力し、前記合成音声と前記録音音声とを組み合わせることを特徴とする音声合成方法。

【請求項 2】 テキスト形式で文章を入力するテキストデータの入力手段と、前記テキストデータから録音音声のテキストデータの切り分けを行うテキストデータ切り分け手段と、前記テキストデータ切り分け手段が前記録音音声のテキストデータの切り分けを行ったとき予め設定した録音音声データベースより録音音声を読み出す録音音声読み出し手段と、前記テキストデータ切り分け手段が前記録音音声のテキストデータの切り分けを行なわなかったとき波形重畳法によって前記テキストデータの合成音声を作成する合成音声作成手段と、前記録音音声を録音した環境と同じ残響特性を前記合成音声に重畳する残響付与手段とを備え、前記合成音声と前記録音音声とを組み合わせることを特徴とする音声合成装置。

【請求項 3】 前記残響特性を付与するとき、前記録音音声を再生する場所の騒音を計測し、計測した騒音レベルと前記録音音声の振幅情報によって前記合成音声の振幅を調整することを特徴とする請求項 1 に記載の音声合成方法。

【請求項 4】 前記録音音声を再生する場所の騒音を計測する騒音計測手段と、前記騒音計測手段によって計測された騒音レベルと、前記録音音声の振幅情報とから前記合成音声の振幅を調整する振幅調整手段とを備えたことを特徴とする請求項 2 に記載の音声合成装置。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、合成音声と録音音声とを組み合わせる再生する音声合成方法および

その装置に関し、特に、両者の音質の差により聴者に与える違和感を低減する音声合成方法およびその装置に関する。

【 0 0 0 2 】

【従来の技術】

従来、複数の異なる音声合成システムを用いた音声合成方法は、例えば特開平 5 - 2 7 7 8 9 号公報に記載されているものがあり、合成音声と録音音声との組み合わせによって再生を行っていた。

【 0 0 0 3 】

【発明が解決しようとする課題】

しかしながら、従来の波形重畳法で作成した合成音声と録音音声を組み合わせることによって音声を作成する音声合成方法は、波形重畳法によって作成された合成音声に録音音声のような録音環境にあった残響がない。

【 0 0 0 4 】

従って、合成音声と録音音声とを組み合わせた音声を聴者に聞かせると、合成音声と録音音声とのつなぎ部分で、音質の差により聴者に違和感を与えてしまうという問題があった。

【 0 0 0 5 】

本発明は、このような従来の問題を解決するためになされたもので、合成音声と録音音声を組み合わせて音声を再生する場合、音質の差により聴者に与える違和感を低減する音声合成方法およびその装置を提供するものである。

【 0 0 0 6 】

【課題を解決するための手段】

本発明の音声合成方法は、テキストデータを録音音声で再生するときには、前記録音音声をそのまま出力し、波形重畳法で作成した合成音声で出力するときには、前記録音音声を録音した環境と同じ残響特性をもつ残響特性を付与して出力し、前記合成音声と前記録音音声とを組み合わせ再生することとした。

【 0 0 0 7 】

この方法により、録音音声と合成音声に同じ残響特性を持たせることができ、録音音声と合成音声とを組み合わせ再生する場合、合成音声と録音音声とのつ

な部分で音質の差により聴者が抱く違和感を低減させることができる。

【 0 0 0 8 】

また、本発明の音声合成装置は、テキスト形式で文章を入力するテキストデータの入力手段と、前記テキストデータから録音音声のテキストデータの切り分けを行うテキストデータ切り分け手段と、前記テキストデータ切り分け手段が前記録音音声のテキストデータの切り分けを行ったとき予め設定した録音音声データベースより録音音声を読み出す録音音声読み出し手段と、前記テキストデータ切り分け手段が前記録音音声のテキストデータの切り分けを行なわなかったとき波形重畳法によって前記テキストデータの合成音声を作成する合成音声作成手段と、前記録音声を録音した環境と同じ残響特性を前記合成音声に重畳する残響付与手段とを備え、前記合成音声と前記録音声をとを組み合わせ再生することとした。

【 0 0 0 9 】

この構成により、録音音声と合成音声に同じ残響特性を持たせることができ、録音音声と合成音声とを組み合わせ再生する場合、合成音声と録音音声とのつながり部分で音質の差により聴者が抱く違和感を低減させる装置とすることができる。

【 0 0 1 0 】

また、前記残響特性を付与するとき、前記録音声を再生する場所の騒音を計測し、計測した騒音レベルと録音音声の振幅情報によって前記合成音声の振幅を調整することとした。

【 0 0 1 1 】

この方法により、録音音声と合成音声に同じ残響特性を持たせることができ、また録音音声に比べて明瞭度の低い合成音声の振幅を大きくすることができる。

【 0 0 1 2 】

従って、録音音声と合成音声とを組み合わせ再生する場合、合成音声と録音音声とのつながり部分で音質の差により聴者が抱く違和感を低減し、騒々しい環境でも合成音声の明瞭度の低下を抑えることができる。

【 0 0 1 3 】

また、前記録音音声を再生する場所の騒音を計測する騒音計測手段と、前記騒音計測手段によって計測された騒音レベルと、前記録音音声の振幅情報とから前記合成音声の振幅を調整する振幅調整手段とを備えることとした。

【0014】

この構成により、録音音声と合成音声に同じ残響特性を持たせることができ、また録音音声に比べて明瞭度の低い合成音声の振幅を大きくすることができる。

【0015】

従って、録音音声と合成音声とを組み合わせ再生する場合、合成音声と録音音声とのつなぎ部分で音質の差により聴者が抱く違和感を低減し、騒々しい環境でも合成音声の明瞭度の低下を抑えることができる。

【0016】

【発明の実施の形態】

以下、本発明の実施の形態について、図面を用いて説明する。

【0017】

図1は、本発明の第1の実施の形態の音声合成方法のフローを示す。

【0018】

図1に示すように、この音声合成方法は、テキスト形式で入力する文章のテキストデータの入力処理101と、入力したテキストデータが固定メッセージ部分を有するかどうかの判断処理102と、録音音声読み出し処理103と、合成音声作成処理104と、残響特性付与処理105と、テキストデータの入力が終了したかの判断処理106とを行う。

【0019】

以上のように処理される音声合成方法について、図1を用いてその動作を説明する。

【0020】

まず、文章の情報出力要求をテキスト形式で入力する（101）。続いて入力したテキストデータの中から、録音音声再生する固定メッセージ部分と、合成音声を作成し再生する部分を切り分ける（102）。もし録音音声再生する固定メッセージ部分なら録音音声を読み出す（103）。そうでなければ、合成音

声を作成し（１０４）、録音音声を録音した環境と同じ残響特性を付与する（１０５）。このようにテキストデータすべての文章が終了するまで一連の処理を繰り返し、すべての文章を音声化する（１０６）。

【００２１】

この発明の第１の実施の形態の音声合成方法によれば、録音音声を録音した環境と同じ残響特性を重畳する残響特性付与処理を設けることにより、録音音声と合成音声に同じ残響特性を持たせることができ、録音音声と合成音声とを組み合わせ再生する場合、合成音声と録音音声とのつなぎ部分で音質の差により聴者に与える違和感を低減させることができる。

【００２２】

図２は、本発明の第２の実施の形態の音声合成装置のブロック図を示す。

【００２３】

図２に示すようにこの音声合成装置は、テキストデータ入力手段２０１と、テキストデータ切り分け手段２０２と、録音音声テキストデータベース２０３と、合成音声作成手段２０４と、音声素片データベース２０５と、残響特性付与手段２０６ａと、残響特性付与手段２０６ａを有するポストフィルタ２０６と、録音音声データベース２０７と、録音音声読み出し手段２０８と、音声出力手段２０９とにより構成されている。

【００２４】

以上のように構成された音声合成装置について、図２を用いてその動作を説明する。

【００２５】

まず、テキストデータをテキストデータ入力手段２０１によって入手する。続いてテキストデータ入力手段２０１によって入手したテキストデータと、録音音声テキストデータベース２０３に保存してある録音音声のテキストデータとを比較することにより、テキストデータ切り分け手段２０２においてテキストデータの中から、録音音声のテキストデータベース２０３にある固定メッセージ部分のテキストデータとそうでない部分のテキストデータを切り分ける。

【００２６】

テキストデータ切り分け手段 2 0 2 によりテキストデータが切り分けられたとき、テキストデータが固定メッセージ部分にある場合には、録音音声データベース 2 0 7 から録音音声読み出し手段 2 0 8 によって録音音声を読み込む。

【 0 0 2 7 】

テキストデータが固定メッセージ部分にない場合には、合成音声作成手段 2 0 4 で音声素片データベース 2 0 5 を使い波形重畳法によって合成音声を作成し、ポストフィルタ 2 0 6 の残響特性付与手段 2 0 6 a によって録音音声を録音した環境と同じ残響を重畳する。最後にすべての文章を録音音声と合成音声とを組み合わせ、音声出力手段 2 1 0 から音声として出力する。

【 0 0 2 8 】

この本発明の第 2 の実施の形態の音声合成装置によれば、録音音声を録音した環境と同じ残響特性を重畳する残響特性付与手段を設けることにより、録音音声と合成音声に同じ残響特性を持たせることができ、録音音声と合成音声とを組み合わせ再生する場合、合成音声と録音音声とのつなぎ部分で音質の差により聴者に与える違和感を低減させることができる。

【 0 0 2 9 】

図 3 は、本発明の第 3 の実施の形態の音声合成方法のフローを示す。

【 0 0 3 0 】

図 3 に示すように、この音声合成方法は、テキストデータ入力処理 1 0 1 と、入力したテキストデータが固定メッセージ部分かどうかの判断処理 1 0 2 と、録音音声読み出し処理 1 0 3 と、合成音声作成処理 1 0 4 と、残響特性付与処理 1 0 5 と、テキストデータの入力が終了したかの判断処理 1 0 6 と、振幅調整処理 1 0 7 とを行う。

【 0 0 3 1 】

以上のように処理を行う音声合成方法について、図 3 を用いてその動作を説明する。

【 0 0 3 2 】

まず、文章の情報出力要求をテキスト形式で入力する（1 0 1）。続いて入力したテキストデータの中から、録音音声を再生する固定メッセージ部分と、合成

音声を作成し再生する部分を切り分ける（102）。

【0033】

もし、テキストデータが録音音声を再生する固定メッセージ部分なら録音音声を読み込む（103）。そうでなければ、合成音声を作成し（104）、録音音声を録音した環境と同じ残響特性を付与する（105）。

【0034】

さらに、残響特性を付与した後、音声を再生する場所の騒音騒音レベルと録音音声の振幅情報によって合成音声の振幅を調整する（107）。このようにテキストデータのすべての文章が終了するまで一連の処理を繰り返し、すべての文章を音声化する（106）。

【0035】

この本発明の第3の実施の形態の音声合成方法によれば、録音音声を録音した環境と同じ残響特性を重畳する残響特性付与処理と、音声を再生する場所の騒音騒音レベルと録音音声の振幅情報をもとに合成音声の振幅を調整する振幅調整処理とを設けることにより、録音音声と合成音声に同じ残響特性を持たせることができ、また録音音声に比べて明瞭度の低い合成音声の振幅を大きくすることができる。

【0036】

従って、録音音声と合成音声とを組み合わせ再生する場合、合成音声と録音音声とのつなぎ部分で音質の差により聴者に与える違和感を低減させることができ、さらに騒音による合成音声の明瞭度の低下を抑えることができる。

【0037】

図4は、本発明の第4の実施の形態の音声合成装置のブロック図を示す。

【0038】

図4に示すように、この音声合成装置は、テキストデータ入力手段201と、テキストデータ切り分け手段202と、録音音声テキストデータベース203と、合成音声作成手段204と、音声素片データベース205と、残響特性付与手段206aと、振幅調整手段206bと、残響特性付与手段206aと振幅調整手段206bとを有するポストフィルタ206と、録音音声データベース207

と、録音音声読み出し手段 2 0 8 と、音声出力手段 2 0 9 と、騒音計測手段 2 1 0 とにより構成されている。

【 0 0 3 9 】

以上のように構成された音声合成装置について、図 4 を用いてその動作を説明する。

【 0 0 4 0 】

まず、テキストデータをテキストデータ入力手段 2 0 1 によって入手する。続いてテキストデータ入力手段 2 0 1 によって入手したテキストデータと、録音音声テキストデータベース 2 0 3 に保存している録音した音声のテキストデータとを比較することにより、テキストデータ切り分け手段 2 0 2 においてテキストデータの中から、録音音声のテキストデータベース 2 0 3 にある固定メッセージ部分のテキストデータとそうでない部分のテキストデータを切り分ける。

【 0 0 4 1 】

テキストデータ切り分け手段 2 0 2 によりテキストデータが切り分けられたとき、テキストデータが固定メッセージ部分にある場合には、録音音声データベース 2 0 7 から録音音声読み出し手段 2 0 8 によって録音音声を読み込む。

【 0 0 4 2 】

テキストデータが固定メッセージ部分にない場合には、合成音声作成手段 2 0 4 で音声素片データベース 2 0 5 を使い波形重畳法によって合成音声を作成し、ポストフィルタの残響特性付与手段 2 0 6 a によって録音音声を録音した環境と同じ残響を重畳する。さらに、騒音計測手段 2 1 0 で音声を再生する場所の騒音レベルを計測し、前記計測騒音レベルと録音音声読み出し手段 2 0 8 からの録音音声の振幅情報をもとに、振幅調整手段 2 0 6 a で合成音声の振幅を調整する。最後に、すべての文章を録音音声と合成音声とを組み合わせ、音声出力手段 2 0 9 から音声として出力する。

【 0 0 4 3 】

この本発明の第 4 の実施の形態の音声合成装置によれば、録音音声を録音した環境と同じ残響特性を重畳する残響特性付与手段と、音声を再生する場所の騒音騒音レベルと録音音声の振幅情報をもとに合成音声の振幅を調整する振幅調整手

段とを設けることにより、録音音声と合成音声に同じ残響特性を持たせることができ、また録音音声に比べて明瞭度の低い合成音声の振幅を大きくすることができる。

【 0 0 4 4 】

従って、録音音声と合成音声とを組み合わせ再生する場合、合成音声と録音音声とのつなぎ部分で音質の差により聴者に与える違和感を低減することができ、さらに騒音による合成音声の明瞭度の低下を抑えることができる。

【 0 0 4 5 】

【発明の効果】

本発明の音声合成方法は、テキストデータを録音音声で再生するときには、前記録音音声のまま出力し、波形重畳法で作成した合成音声で出力するときには、前記録音声を録音した環境と同じ残響特性をもつ残響特性を付与して出力し、前記合成音声と前記録音声を組み合わせ再生することとした。

【 0 0 4 6 】

この方法により、録音音声と合成音声に同じ残響特性を持たせることができ、録音音声と合成音声とを組み合わせ再生する場合、合成音声と録音音声とのつなぎ部分で音質の差により聴者が抱く違和感を低減させることができる。

【図面の簡単な説明】

【図 1】

本発明の第 1 の実施の形態の音声合成方法のフローを示す図

【図 2】

本発明の第 2 の実施の形態の音声合成装置のブロック図

【図 3】

本発明の第 3 の実施の形態の音声合成方法のフローを示す図

【図 4】

本発明の第 4 の実施の形態の音声合成装置のブロック図

【符号の説明】

2 0 1 テキストデータ入力手段

2 0 2 テキストデータ切り分け手段

2 0 3 録音音声テキストデータベース

2 0 4 合成音声作成手段

2 0 5 音声素片データベース

2 0 6 ポストフィルタ

2 0 6 a 残響特性付与手段

2 0 6 b 振幅調整手段

2 0 7 録音音声データベース

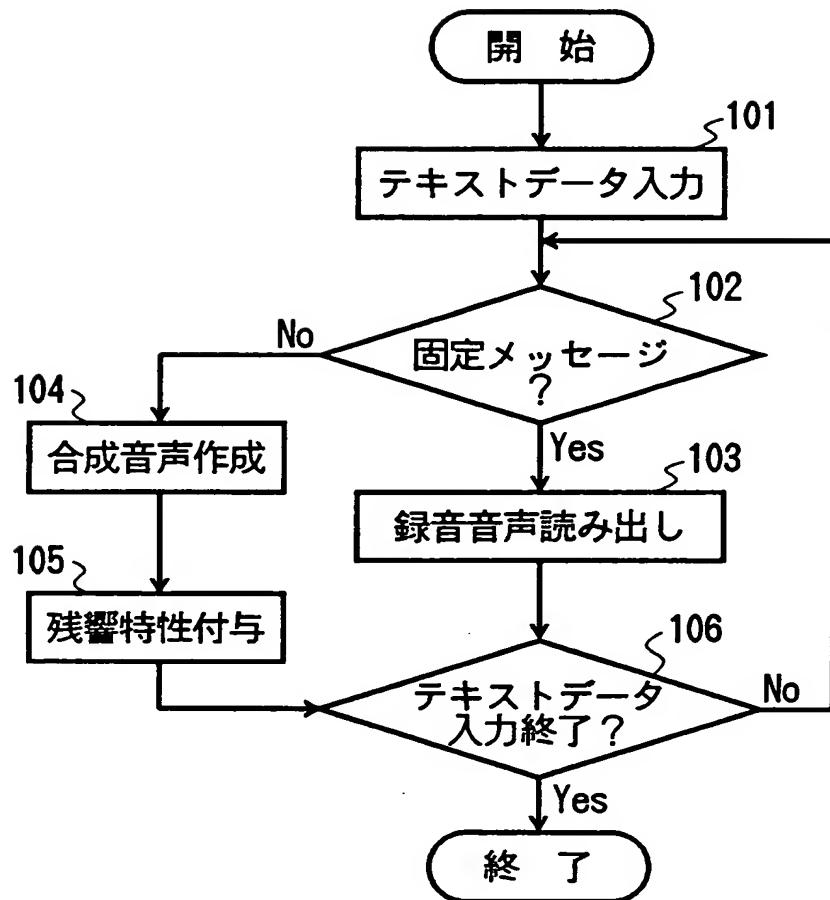
2 0 8 録音音声読み出し手段

2 0 9 音声出力手段

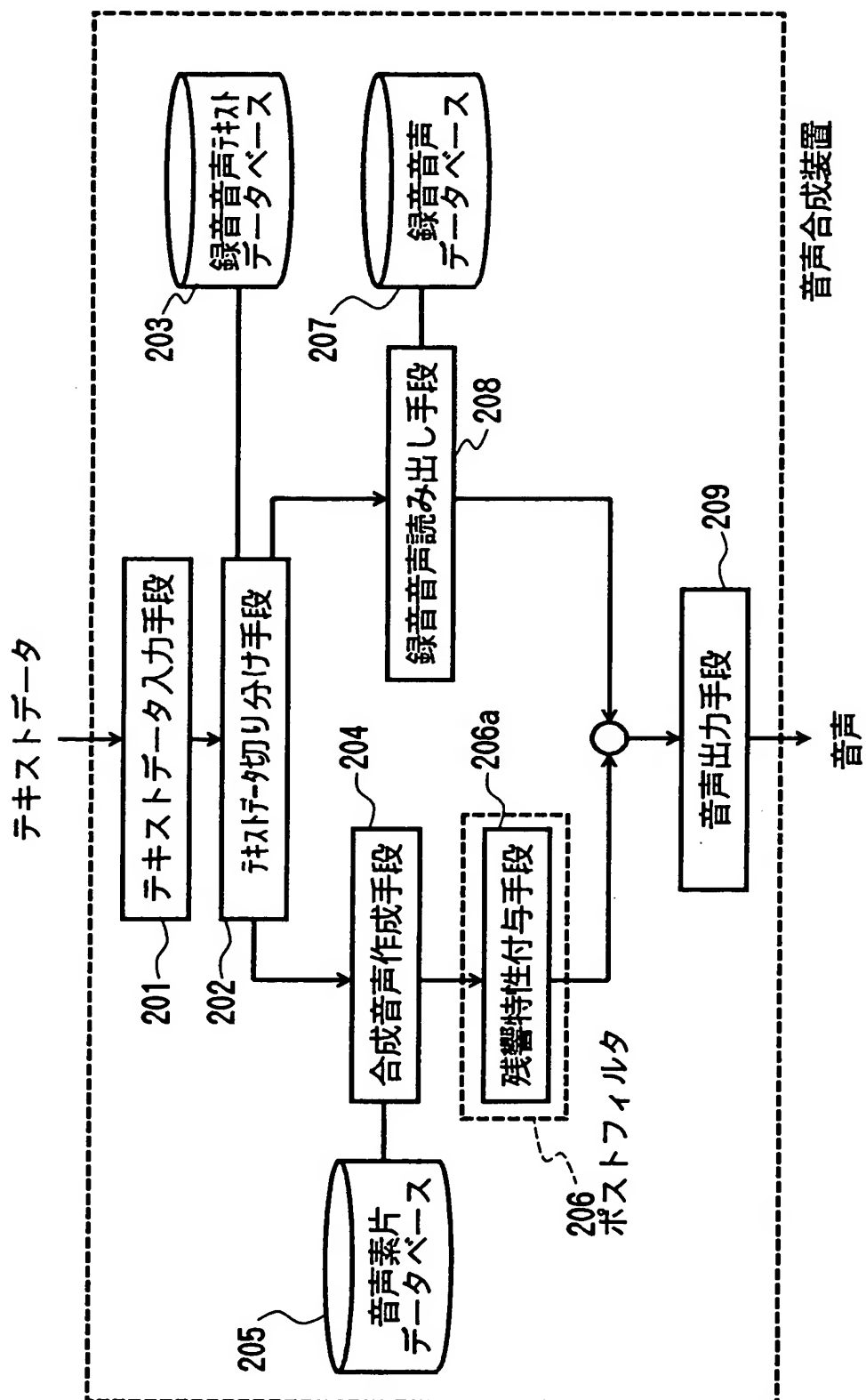
2 1 0 騒音計測手段

【書類名】 図面

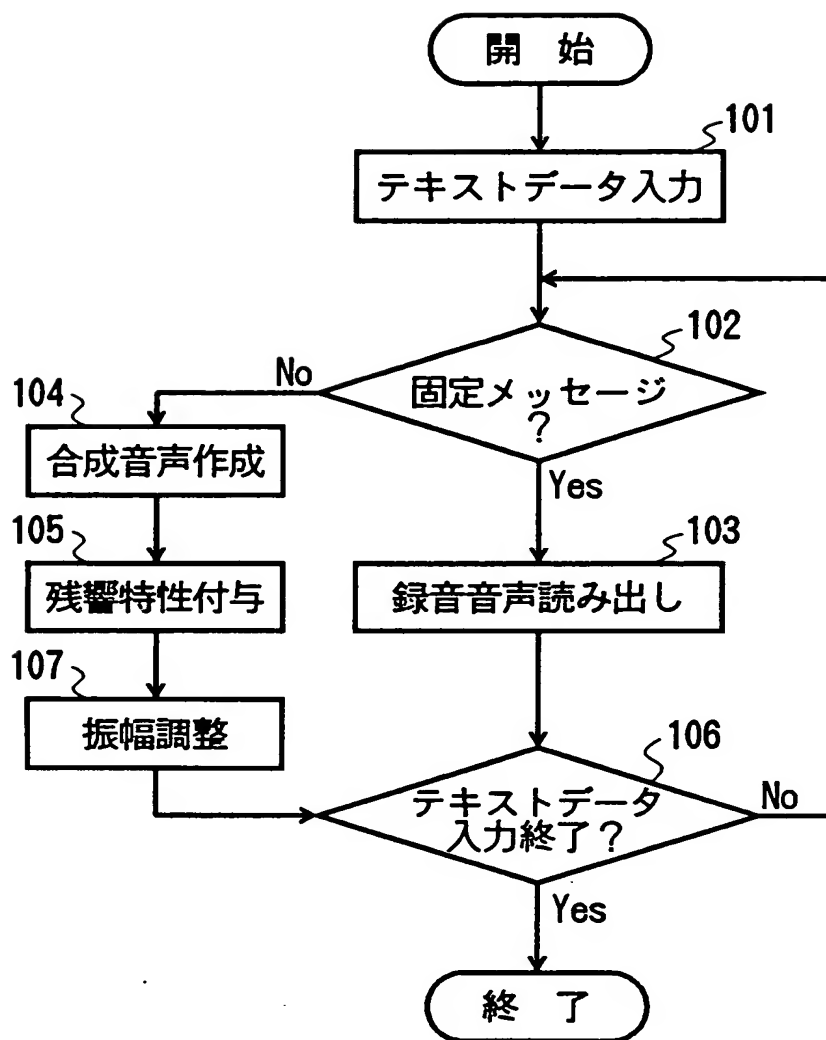
【図 1】



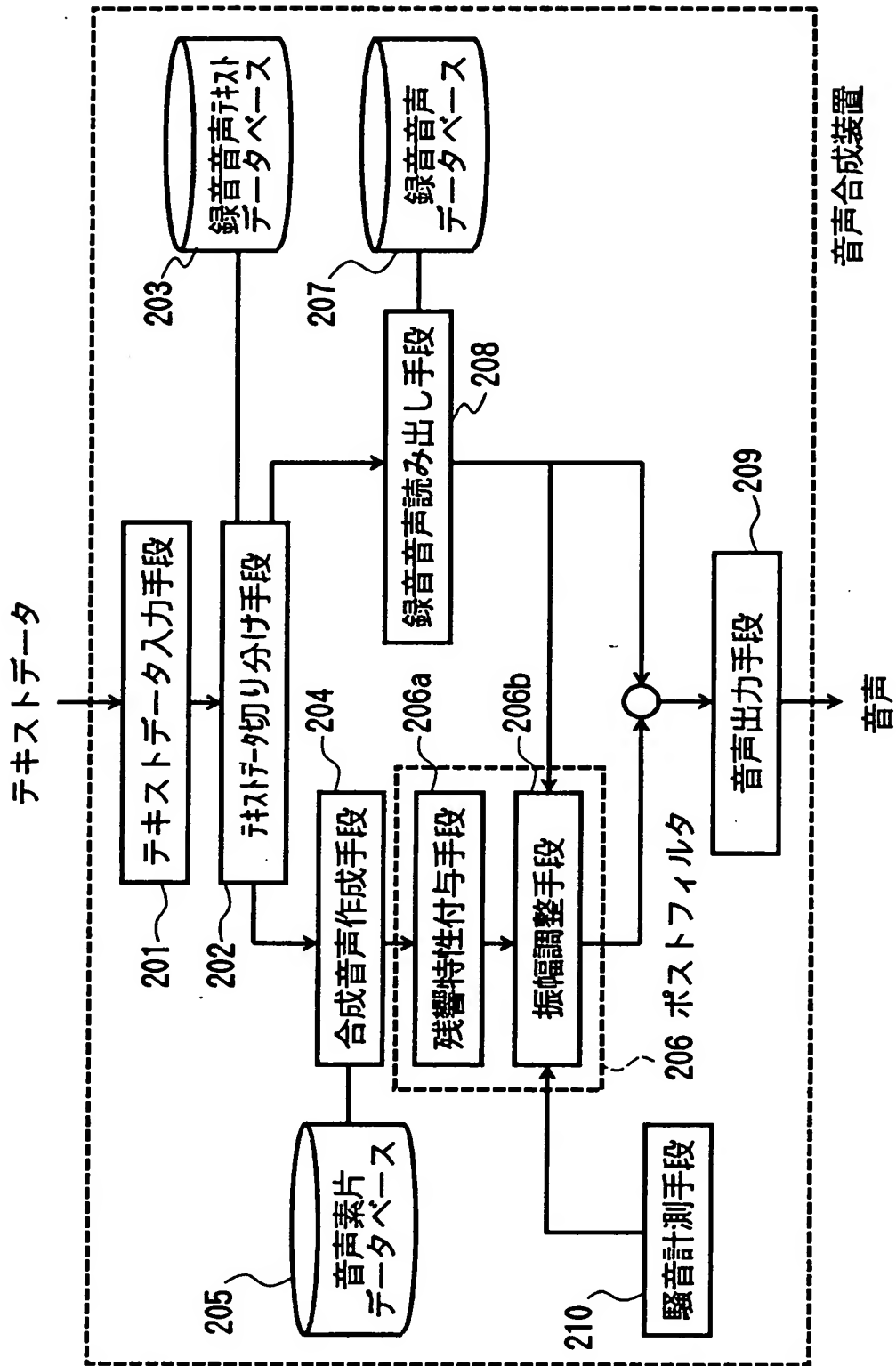
【図 2】



【図 3】



【図 4】



【書類名】 要約書

【要約】

【課題】 合成音声と録音音声を組み合わせて音声を再生する場合、音質の差により聴者に与える違和感を低減する音声合成方法およびその装置を提供すること。

【解決の手段】 本発明の音声合成方法は、テキストデータを録音音声で再生するときには、前記録音音声をそのまま出力し、波形重畳法で作成した合成音声で出力するときには、前記録音声を録音した環境と同じ残響特性をもつ残響特性を付与して出力し、前記合成音声と前記録音声を組み合わせて再生することとした。

【選択図】 図1

出 願 人 履 歷 情 報

識別番号 [000005821]

1. 変更年月日	1990年 8月28日
[変更理由]	新規登録
住 所	大阪府門真市大字門真1006番地
氏 名	松下電器産業株式会社